Statistical mechanics of isomerization dynamics in liquids and the transition state approximation[a]

David Chandler[b]

School of Chemical Sciences, University of Illinois, Urbana, Illinois 61801
and Department of Chemistry, Columbia University, New York, New York 10027

(Received 17 October 1977)

In this article, time correlation function methods are used to discuss classical isomerization reactions of small nonrigid molecules in liquid solvents. Molecular expressions are derived for a macroscopic phenomenological rate constant. The form of several of these equations depend upon what ensemble is used when performing averages over initial conditions. All of these formulas, however, reduce to one final physical expression whose value is manifestly independent of ensemble. The validity of the physical expression hinges on a separation of time scales and the plateau value problem. The approximations needed to obtain transition state theory are described and the errors involved are estimated. The coupling of the reaction coordinate to the liquid medium provides the dissipation necessary for the existence of a plateau value for the rate constant, but it also leads to failures of Wigner's fundamental assumption for transition state theory. We predict that for many isomerization reactions, the transmission coefficient will differ significantly from unity and that the difference will be a strong function of the thermodynamic state of the liquid solvent.

I. INTRODUCTION

In a recent series of articles,[1,2] we have studied the equilibrium statistical mechanics of small nonrigid molecules. We derived microscopic expressions which describe how condensed phase environments alter the average conformational structures of molecules from what is found in the gas phase.

The *trans–gauche* equilibrium of n-butane is perhaps the simplest example of this phenomenon (see Fig. 1). The dihedral angle ϕ is the only appreciable intramolecular degree of freedom. The intramolecular energetics involved when rearranging the conformation of the molecule is denoted by $V(\phi)$. In the gas phase, the distribution function for the variable ϕ is

$$s^{(0)}(\phi)=e^{-\beta V(\phi)}\Big/\int_{-\pi}^{\pi}d\phi\, e^{-\beta V(\phi)}, \tag{1.1}$$

where the superscripts 0 indicates a gas phase result, and $\beta^{-1}=k_BT$.

Since conformational rearrangements in liquids involve the displacement of solvent molecules, the ability of the solvent to literally digest a particular conformational species must be accounted for. The distribution function for the nonrigid solute in a liquid is similar to $s^{(0)}(\phi)$ in Eq. (1.1); but the bare intramolecular potential $V(\phi)$ must be replaced with a free energy function. Indeed, the exact result is[1,2]

$$e^{-\beta V(\phi)-\beta w(\phi)}\Big/\int_{-\pi}^{\pi}d\phi\, e^{-\beta V(\phi)-\beta w(\phi)}, \tag{1.2}$$

where $w(\phi)$ is the change in Helmholtz free energy of the liquid solvent due to rotating the dihedral angle of a single solute molecule from zero (our choice of the state of zero energy) to ϕ. The quantity $\exp[-\beta w(\phi)]$ is proportional to the cavity distribution $y^*(\phi)$ for the solute.[1] Thus, Eq. (1.2) can also be written as

$$s(\phi)=s^{(0)}(\phi)y^*(\phi)\Big/\int_{-\pi}^{\pi}d\phi\, s^{(0)}(\phi)y^*(\phi). \tag{1.2$'$}$$

From Eq. (1.2′) and the rotational isomeric model,[3] it follows that the mole-fraction equilibrium constant for *trans–gauche* equilibrium obeys the equation

$$(x_g/x_t)=(x_g/x_t)^{(0)}(y_g^*/y_t^*). \tag{1.3}$$

Simple numerical calculations[1,4] of the cavity distribu-

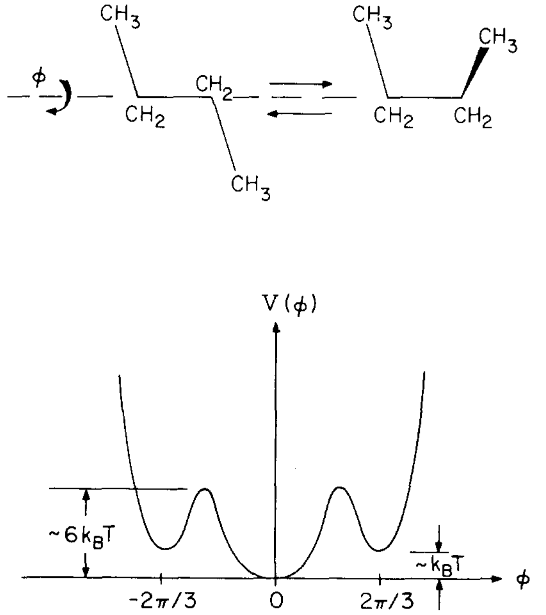

FIG. 1. *Trans–gauche* isomerization for n-butane. ϕ is the dihedral angle. The *trans* conformer is shown schematically on the left. ($\phi\approx 0$ for that state). The *gauche* conformers, pictured on the right, are located at $\phi\approx\pm 2\pi/3$. $V(\phi)$ is the intramolecular potential energy for n-butane. The graph of $V(\phi)$ is schematic.

---

[a] This research has been suppported by the National Science Foundation and by the donors of the Petroleum Research Fund as administered by the American Chemical Society.

[b] On leave at Columbia University from the University of Illinois for the 1977–78 academic year.

J. Chem. Phys. 68(6), 15 Mar. 1978 0021-9606/78/6806-2959$01.00

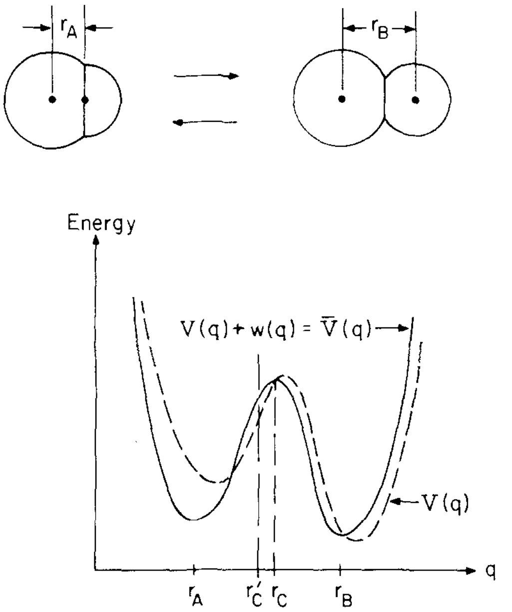

FIG. 2. The two-state one-dimensional isomerization model discussed in the text. One can imagine the hypothetical molecule drawn in the figure. The potential of mean force $\bar{V}(q)$ is compared with the bare intramolecular isomerization potential $V(q)$ in the graph. The points r_C and r_C' are two alternatives for the point which divides state A from state B.

tions for n-butane in the *gauche* and *trans* states y_g^* and y_t^*, respectively, show that (y_g^*/y_t^*) differs significantly from unity (the ideal gas result), when n-butane is immersed in a dense fluid such as liquid carbon tetrachloride. Thus, liquid phase values of (x_g/x_t) differ substantially from the gas phase ratio $(x_g/x_t)^{(0)}$.

These remarks refer to static properties. It is reasonable to expect the dynamics of isomerization equilibrium is also affected by condensed phases. Indeed, according to classical transition state theory,[5-7] the rate constant for going from *trans* to *gauche* states, $k_{t\to g}$, should be related to its value in the gas phase at the same temperature by the equation

$$k_{t\to g}=k_{t\to g}^{(0)}\,\frac{y_\ddagger^*}{y_t^*}\,, \tag{1.4}$$

where $y_\ddagger^*$ denotes the value of the cavity distribution function for the activated (transition) state. Presumably that state lies close to the value of ϕ at which $V(\phi)$ is a maximum between the trans ($\phi=0$) and gauche ($\phi=\pm 2\pi/3$) states.

In this article, we explore with time correlation function formalism the conditions for which transition state equations like Eq. (1.4) are valid. There are really three basic questions which must be resolved. The first has to do with the existence of the rate constant. It must be possible to write a rate law; and this can be done only if the time scale for isomerizations is much larger than the times which characterize the dynamics of all the other degrees of freedom that couple with the conformational coordinates. The second question concerns the correct molecular expression for the rate constant once its existence is accepted. The third asks what dynamical approximations must be made to obtain the transition state theory formula from the molecular expression for the rate constant.

We focus attention on the second and third questions. We assume that a rate law for isomerization reactions is established *experimentally*. With this assumption as a starting point, there is still a lot to be done. Our analysis touches upon ideas that have appeared in the literature for many years. For example, the correlation function expressions for rate constants which are derived in Secs. III, V, and VII are closely related to formulas presented by Yamamoto[8] and Zwanzig[9] and more recently by Kutz, Oppenheim, and Ben-Reuven[10] and by Stillinger.[11] The physical picture discussed in Sec. VIII with which the transition state theory approximation is justified from a correlation function expression is similar to the so-called[6] "fundamental assumption" of Wigner.[5] A major purpose of this article is to summarize these old ideas in a common language. The results presented herein that we believe are new involve nuances associated with equilibrium statistical mechanics of chemical equilibrium in condensed phases,[1,2] with the ensemble dependence of correlation function formulas,[12,13] and with the plateau value problem.[14-17] These subtleties are addressed in both the main text and the Appendix.

II. SIMPLE TWO STATE SYSTEM AND PHENOMENOLOGICAL DESCRIPTION OF ISOMERIZATION DYNAMICS

For notational convenience, consider a hypothetical isomerization process that is even simpler than the *trans–gauche* transitions in n-butane. In particular, imagine a molecule with a single internal degree of freedom, described by the coordinate q, and the intramolecular potential $V(q)$ drawn in Fig. 2. $V(q)$ contains two minima. One is located at $q=r_A$. When q is close to that point, the molecule is in state A. Similarly, the region near r_B is associated with state B. The dividing "surface" (i.e., point) between the two states is the position r_C. This position r_C should be located near the maximum in $V(q)$ which separates the two states.

We find that exact relations are insensitive to the precise choice of r_C. However, to derive transition state theory expressions—which are approximate—the choice of r_C is crucial. The reader should note that the location of the extremum in $V(q)$ is not necessarily the same as that for the potential of mean force for the internal variable q, namely, $\bar{V}(q)$.[1,2] As we have already mentioned, the solvent structure can significantly alter the effective intramolecular potential from its nature in the gas phase. Qualitative differences between $V(q)$ and $\bar{V}(q)$ are shown schematically in Fig. 2. In that figure, the position r_C is taken as the location of the maximum for $\bar{V}(q)$ because that choice is useful in obtaining transition state theory. This matter will be discussed in Sec. VIII.

With r_C defined as the dividing surface between states A and B, we can define the numbers of molecules in states A and B as follows: Let $q_i(t)$ denote the internal coordinate for the ith solute molecule at time t. There are N solute molecules in the system. The characteristic functions for states A and B are

$$H_A(r)=\begin{cases}1, & r<r_C\,,\\ 0, & r>r_C\,,\end{cases} \tag{2.1a}$$

and

$$H_B(r)=1-H_A(r)=\begin{cases}0, & r<r_C\,,\\ 1, & r>r_C\,.\end{cases} \tag{2.1b}$$

Thus, the number of molecules in state A at time t is

$$N_A(t)=\sum_{i=1}^{N} H_A[q_i(t)]\,, \tag{2.2a}$$

and similarly

$$N_B(t)=N-N_A(t)=\sum_{i=1}^{N} H_B[q_i(t)]\,. \tag{2.2b}$$

For an open system, the total number of solute molecules N is a function of time since particles can flow in and out of the system. However, for a closed system N is a constant of the motion. This simple difference between open and closed systems suggests that differences must exist between grand canonical and canonical ensemble analyses of isomerization reactions. Indeed, the two ensembles yield different equations at nearly all the intermediate steps. Only the final physical expressions are the same. In the main text that follows, we study a closed system which is characterized by the canonical ensemble. The open system (grand canonical ensemble) analysis is studied in the Appendix.

The average numbers of A and B isomers are $\langle N_A\rangle$ and $\langle N_B\rangle$, where the pointed brackets denote the equilibrium ensemble average. The equilibrium constant is

$$K=\langle N_B\rangle/\langle N_A\rangle=x_B/x_A\,, \tag{2.3}$$

where x_B and x_A denote the average mole fractions of B and A isomers, respectively. The application of a disturbance changes the observed numbers from their equilibrium values $\langle N_A\rangle$ and $\langle N_B\rangle$ to time dependent nonequilibrium values $\langle N_A\rangle_{\rm ne}(t)$ and $\langle N_B\rangle_{\rm ne}(t)$. At the macroscopic level, the time dependence of these numbers can be characterized by a rate law.

We shall assume that at low solute concentrations, the experimentalist finds a simple linear rate law

$$\frac{d\langle N_A\rangle_{\rm ne}(t)}{dt}=-k_{A\to B}\langle N_A\rangle_{\rm ne}(t)+k_{B\to A}\langle N_B\rangle_{\rm ne}(t)\,. \tag{2.4}$$

Notice that there are no diffusion terms in Eq. (2.4). Hence, Eq. (2.4) can be correct only for a closed system where $N=N_A(t)+N_B(t)=$constant. The solution of Eq. (2.4) is

$$\langle N_A\rangle_{\rm ne}(t)=\langle N_A\rangle+[\langle N_A\rangle_{\rm ne}(0)-\langle N_A\rangle]e^{-t/\tau_{\rm rxn}}\,, \tag{2.5}$$

where

$$\tau_{\rm rxn}^{-1}=k_{A\to B}+k_{B\to A}\,. \tag{2.6}$$

The two rate constants are related by the detailed balance condition,

$$K=k_{A\to B}/k_{B\to A}\,, \tag{2.7}$$

so that an alternative formula for the reaction relaxation time $\tau_{\rm rxn}$ is

$$\tau_{\rm rxn}^{-1}=k_{A\to B}/x_B\,. \tag{2.8}$$

It is probably impossible to *derive* a macroscopic rate equation like Eq. (2.4) from realistic microscopic laws. However, Eq. (2.4) is a reasonable model of what would be found experimentally for a closed system. Some phenomenological expression is required as a starting point for a correlation function analysis of a rate constant—or any other transport coefficient. The precise identification of microscopic quantities with the macroscopic rate constants depends upon the particular form of the phenomenological equation. If experimentalists inform us of another rate equation, it too can be analyzed in an identical fashion.

III. FIRST CORRELATION FUNCTION EXPRESSION

We now discuss our first microscopic equation for the rate constants in Eq. (2.4). An equivalent memory function expression is discussed in Sec. V. Then a related approximate though physical expression is derived in Secs. VI and VII. The approximate equation is our "second correlation function expression."

Consider the *equilibrium* time correlation function

$$C(t)=\langle\delta N_A\delta N_A(t)\rangle\,, \tag{3.1}$$

where

$$\delta N_A(t)=N_A(t)-\langle N_A\rangle\,, \tag{3.2}$$

$$\delta N_A=\delta N_A(0)\,, \tag{3.3}$$

and the equilibrium ensemble average is over initial conditions (all the coordinates and momenta of *all* the particles in the system, both solute and solvent molecules). The behavior of $C(t)$ is linked to $\langle N_A\rangle_{\rm ne}(t)$ in the linear regime by the fluctuation–dissipation theorem.[18] In particular, close to equilibrium

$$\frac{\tilde C(s)}{C(0)}=\int_0^\infty dt\, e^{-st}\left(\frac{\langle N_A\rangle_{\rm ne}(t)-\langle N_A\rangle}{\langle N_A\rangle_{\rm ne}(0)-\langle N_A\rangle}\right)\,, \tag{3.4}$$

where $\tilde C(s)$ denotes the Laplace transform of $C(t)$. The phenomenological result, Eq. (2.5), can be used at small values of the Laplace transform variable s. Provided the rate law is correct,

$$[\tilde C(s)/C(0)]\sim(s+\tau_{\rm rxn}^{-1})^{-1}\,, \tag{3.5}$$

for small s. As a result,

$$\tau_{\rm rxn}=\int_0^\infty dt\,\frac{\langle\delta N_A\delta N_A(t)\rangle}{\langle(\delta N_A)^2\rangle}\,, \tag{3.6}$$

which provides an exact connection between the phenomenological rate constant and a microscopic canonical ensemble time correlation function. Of course, this formula is not a new result. See, for example, Ref. 10. An alternative formula for $\tau_{\rm rxn}$ has been derived by Stillinger.[11] His result (in our notation) is

$$2\tau_{\rm rxn}\langle(\delta N_A)^2\rangle=-\int_0^\infty dt\, t^2\,\frac{d^2C(t)}{dt^2}\,,$$

with

$$\langle(\delta N_A)^2\rangle = \int_0^\infty dt\, t\, \frac{d^2C(t)}{dt^2} ,$$

which appears different than Eq. (3.6). However, one may integrate by parts twice to show that Stillinger's formula is equivalent to Eq. (3.6).

A remarkable aspect of Eq. (3.6) is that it is a canonical ensemble result. It is not valid when the grand canonical ensemble is used to perform the average over initial conditions. This situation is similar to that found by Green[13] when he considered the differences between canonical and microcanonical ensemble calculations of the thermal conductivity. It has long been appreciated that apparent discrepancies of this sort occur because the second moments of fluctuating quantities depend on the ensemble used.[9] When proper allowance is made for the effects of changing ensembles, the resulting formulas should yield identical values (in the thermodynamic limit). The grand canonical formula that corresponds to Eq. (3.6) is discussed in the Appendix. We find that provided it is meaningful to discuss a rate constant, the actual values for $\tau_{\rm rxn}$ do agree in the canonical and grand canonical ensembles even though the fluctuation formulas appear different.

IV. PROPERTIES OF $C(t)$

To proceed with the correlation function analysis of the rate constant, it is useful to understand certain properties of $C(t)$. The behavior of $C(t)$ near $t=0$ is particularly significant. For a closed system, the fluctuations of N_A are coupled to those of N_B due to the constraint $N_A + N_B = N =$ constant. This constraint leads to the canonical ensemble result[19]

$$C(0) = \langle(\delta N_A)^2\rangle = \langle N_A\rangle\langle N_B\rangle/N , \qquad (4.1)$$

which is valid for low solute concentrations. [The open system (grand canonical) result is $\langle(\delta N_A)^2\rangle = \langle N_A\rangle$. See Appendix A.]

The first derivative of $C(t)$ is

$$\frac{dC(t)}{dt} \equiv \dot{C}(t) = \langle\delta N_A\, \dot{N}_A(t)\rangle = -\langle\dot{N}_A(-t)\delta N_A\rangle$$
$$= -\langle\dot{N}_A \delta N_A(t)\rangle . \qquad (4.2)$$

From Eqs. (2.1) and (2.2) one finds

$$\dot{N}_A(t) = -\sum_{i=1}^{N} \dot{q}_i(t)\delta[r_C - q_i(t)] . \qquad (4.3)$$

Hence, at low solute concentrations, where different isomerization molecules are uncorrelated,

$$\dot{C}(t) = \langle N\dot{q}_1\delta(r_C - q_1)H_A[q_1(t)]\rangle$$
$$= -\langle NH_A(q_1)\dot{q}_1(t)\delta[r_C - q_1(t)]\rangle . \qquad (4.4)$$

Let us first evaluate $\dot{C}(t)$ at precisely $t=0$. The result is

$$\dot{C}(0) = \tfrac{1}{2}\langle N\dot{q}_1\delta(r_C - q_1)\rangle = 0 , \qquad (4.5)$$

where the second equality follows from the fact that the classical equilibrium distribution is an even function of velocities. However, $\dot{C}(t)$ is not continuous near $t=0$. At times arbitrarily close but not precisely at $t=0$, $\dot{C}(t)$ is not zero. To demonstrate this behavior, introduce the infinitesimal positive time ϵ, and consider $\dot{q}_1\delta(r_C - q_1)H_A[q_1(\epsilon)]$. The delta function fixes q_1 at r_C at $t=0$. The characteristic function $H_A[q_1(\epsilon)]$ is nonzero only if $q_1(\epsilon) < r_C$. Since initially q_1 is at r_C, and since ϵ is arbitrarily small, $q_1(\epsilon)$ will be less than r_C only if $\dot{q}_1 < 0$. Hence,

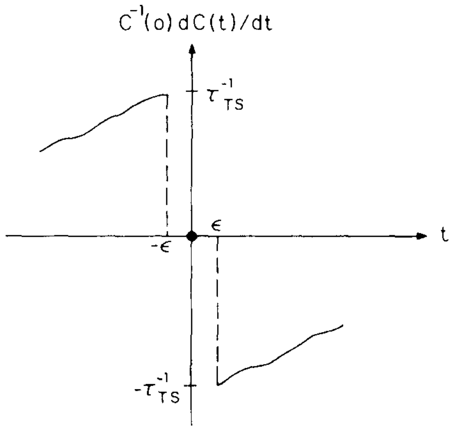

FIG. 3. The short time behavior of $dC(t)/dt$. ϵ is a positive infinitesimal time. See text.

$$\dot{C}(\epsilon) = \langle N\dot{q}_1\delta(r_C - q_1)H_A[q_1(\epsilon)]\rangle$$
$$= \langle N\dot{q}_1\delta(r_C - q_1)\theta(-\dot{q}_1)\rangle$$
$$= -\tfrac{1}{2}\langle N|\dot{q}_1|\delta(r_C - q_1)\rangle , \qquad (4.6)$$

where $\theta(x)$ is the unit step function which is 1 for $x > 0$ and zero otherwise. Similar reasoning demonstrates that $\dot{C}(-\epsilon) = -\dot{C}(\epsilon) \neq 0$.

The right-hand side of Eq. (4.6) is closely related to the transition state theory approximation. That theory gives[5-7]

$$\tau_{\rm rxn}^{-1} \approx \tau_{\rm TS}^{-1} = (2x_Ax_B)^{-1}\langle|\dot{q}_1|\rangle_{r_C} s(r_C) , \qquad (4.7)$$

where

$$s(r_C) = \langle\delta(r_C - q_1)\rangle$$
$$= e^{-\beta\bar{V}(r_C)} \Big/ \int dr\, e^{-\beta\bar{V}(r)} \qquad (4.8)$$

is the probability distribution for finding a molecule with internal coordinate q_1 at position r_C, and $\langle|\dot{q}_1|\rangle_{r_C}$ is the equilibrium ensemble average of $|\dot{q}_1|$ given that q_1 is fixed at r_C. (For a truly one-dimensional problem the specification of that configurational constraint when averaging $|\dot{q}_1|$ is unnecessary since coordinates and their conjugate momenta are uncoupled in the classical equilibrium phase-space distribution function.) By comparing Eqs. (4.1), (4.6), and (4.7), it is seen that

$$[\dot{C}(\epsilon)/C(0)] = -\tau_{\rm TS}^{-1} = -[\dot{C}(-\epsilon)/C(0)] . \qquad (4.9)$$

To summarize, we have shown that the derivative of $C(t)$ is zero at $t=0$, while at $t=\pm\epsilon$, the derivative is $\mp\tau_{\rm TS}^{-1}$. This discontinuous behavior is illustrated in Fig. 3. The behavior implies that the second derivative $\ddot{C}(t)$ contains a delta function at $t=0$. In particular

$$\lim_{\epsilon\to 0+}\int_0^{\epsilon} dt\,[\ddot{C}(t)/C(0)] = -\tau_{\mathrm{TS}}^{-1}. \tag{4.10}$$

If one assumed that $C(t)$ behaved as an exponential, the relaxation time for the exponential decay could be obtained from Eq. (4.9)—the limiting slope as $t\to 0+$. The integral over $C(t)$ yields τ_{rxn}, according to Eq. (3.6). Hence, the exponential assumption would produce $\tau_{\mathrm{rxn}} \approx \tau_{\mathrm{TS}}$. However, such an assumption does not provide a physical interpretation of transition state theory.

V. MEMORY FUNCTION FOR $C(t)$

The inverse relaxation time τ_{rxn}^{-1} plays a role in isomerization dynamics that is analogous to that of transport coefficients in hydrodynamics. This analogy can be pursued by seeking an expression for τ_{rxn}^{-1} that involves a flux autocorrelation function. It will be seen that such an expression cannot hold exactly because there is no experimentally controllable limit in which $N_A(t)$ becomes a conserved variable. Nevertheless, conformational transitions do occur slowly compared to other molecular processes, and something like the traditional hydrodynamics calculations for densities of conserved variables should hold to a good approximation when considering isomerization dynamics. Indeed, much of theoretical chemical kinetics is based upon an expression for τ_{rxn}^{-1} (or equivalently $k_{A\to B}$) that involves an integral of a flux autocorrelation function. To see how such an expression can be derived, we follow the hydrodynamics example, and study the memory function $M(t)$.

The standard Mori–Zwanzig relationships are[20–22]

$$\tilde{C}(s) = [s + \tilde{M}(s)]^{-1}\langle(\delta N_A)^2\rangle, \tag{5.1}$$

with

$$M(t) = \langle(\delta N_A)^2\rangle^{-1}\langle\dot{N}_A e^{QLt}\dot{N}_A\rangle, \tag{5.2}$$

where the projection operator Q is defined by its operation on a general dynamical variable G,

$$QG = (1-P)G, \tag{5.3a}$$

$$PG = \delta N_A\langle(\delta N_A)^2\rangle^{-1}\langle\delta N_A G\rangle, \tag{5.3b}$$

and L is (within a factor of $\sqrt{-1}$) the Liouville operator, so that $LG = \dot{G}$. Note that the particular form of Eqs. (5.1)–(5.3) makes use of the fact that $\langle\delta N_A L N_A\rangle = \langle\delta N_A \dot{N}_A\rangle = 0$. [See Eq. (4.5).] In agreement with the notation of Sec. III, $\tilde{M}(s)$ denotes the Laplace transform of $M(t)$.

The singular behavior of $C(t)$ discussed in Sec. IV implies that $M(t)$ contains a delta function at $t=0$. This feature can be derived by studying the large s behavior of Eq. (5.1). Alternatively it can be verified by studying the time integral of $M(t)$. To carry out the latter procedure, we apply Eq. (5.2) to find

$$\lim_{\epsilon\to 0+}\int_0^{\epsilon} dt\, M(t) = \langle(\delta N_A)^2\rangle^{-1}\lim_{\epsilon\to 0+}\langle\dot{N}_A[N_A(\epsilon) - N_A(0)]\rangle$$
$$= \tau_{\mathrm{TS}}^{-1}, \tag{5.4}$$

where the first equality follows from the fact that $\exp(QL\epsilon)\dot{N}_A = \dot{N}_A(\epsilon) + 0(\epsilon)$, and the second equality is arrived at from the same manipulations carried out in Sec. IV. Since $M(t)$ is integrable, the singularity at $t=0$ does not hinder the applicability of the Mori–Zwanzig formalism.

The zero frequency part of the memory function is related to τ_{rxn}^{-1}. By comparing Eqs. (3.6) and (5.1), one finds

$$\tau_{\mathrm{rxn}}^{-1} = \tilde{M}(0) = \int_0^{\infty} dt\, M(t). \tag{5.5}$$

The delta function part of $M(t)$ contributes τ_{TS}^{-1} to $\tilde{M}(0)$; see Eq. (5.4). Thus,

$$\tau_{\mathrm{rxn}}^{-1} = \tau_{\mathrm{TS}}^{-1} + \int_0^{\infty} dt\,\Delta M(t), \tag{5.6}$$

where $\Delta M(t)$ is the nonsingular part of $M(t)$.

Equations (5.5) and (5.6) are equivalent to Eq. (3.6). They are all exact provided the phenomenological rate law is correct. Notice that the transition state theory approximation is obtained if the integral over $\Delta M(t)$ is negligible compared to τ_{TS}^{-1}. At this stage of our analysis, a physical assessment of this approximation is difficult to obtain because $M(t)$ is not directly related to real dynamical processes. The memory function evolves with a projected propagator $\exp(QLt)$. In the standard hydrodynamics applications of the Mori–Zwanzig formalism, the projected dynamics and real physical processes are connected in the hydrodynamic (long wavelength) limit. A similar connection is *approximately* valid when considering chemical isomerization. To investigate the approximation, it is useful to be aware of the principal time scales involved in isomerization dynamics.

VI. TIME SCALES

The autocorrelation function for $\delta N_A(t)$, $C(t)$, varies on a time scale that is long compared to typical molecular times. Provided $N_A(t)$ is not coupled to another slowly varying variable, $M(t)$ should relax on a short molecular time scale. The projector in Eq. (5.2) guarantees that this will be the case. However, the two preconditions—that $N_A(t)$ is a slow variable, and that $N_A(t)$ does not couple to another slow variable—deserve some comments.

The reason why $N_A(t)$ can be a slow variable is that a potential barrier (or more generally, a free energy barrier) divides state A from state B. Transitions between states A and B are the source for the time dependence of $N_A(t)$. Obviously, as the barrier becomes infinitely high, the transitions become impossible, and $N_A(t)$ becomes a conserved variable. To estimate the time scale over which $N_A(t)$ changes, note that a transition between states A and B implies that at some moment, q_1 was at the transition point r_C. Hence the transition frequency will be proportional to the probability that the molecule is at r_C, $s(r_C)$, times the average flux across the point. Within factors of the order of unity, that frequency is τ_{TS}^{-1}; see Eq. (4.7). Thus, τ_{TS} characterizes the time scale over which $C(t)$ varies.

The correlation functions describing the behavior of other variables which might couple to $N_A(t)$ seem to decay in a time much shorter than τ_{TS}. For example, consider $q_1(t)$ or the coordinates and momenta of neighboring solvent molecules. The characteristic frequency for these variables is $\langle|\dot{q}_1|\rangle(r_C - r_A)^{-1}$, where $(r_C - r_A)$ is chosen as a typical length for the system. (For n-butane, the difference between end-to-end lengths in the trans and gauche states is about 1 Å. The velocity autocorrelation functions for liquids with self-diffusion coefficients in the neighborhood of 10^{-5} cm^2/sec have relaxation times of about $\langle|\dot{q}|\rangle^{-1}$ 1 Å.) Thus, we introduce the "fast" characteristic molecular time

$$\tau_{\mathrm{mol}} = (r_C - r_A)\langle|\dot{q}_1|\rangle^{-1} . \tag{6.1}$$

If our idea that the slow variable $N_A(t)$ couples only to fast variables is correct, then the memory function formalism will lead to a natural perturbation expansion which orders $M(t)$ in powers of the parameter

$$\lambda = \tau_{\mathrm{mol}}/\tau_{\mathrm{TS}} . \tag{6.2}$$

The value of λ is roughly $\exp(-\beta A^\dagger)$, where $A^\dagger$ is the activation free energy, $\bar{V}(r_C) - \bar{V}(r_A)$. To understand this estimate, note that

$$\int dr\, e^{-\beta\bar{V}(r)} \approx (r_C - r_A)(e^{-\beta\bar{V}_A} + e^{-\beta\bar{V}_B}) , \tag{6.3}$$

and

$$x_A \approx e^{-\beta\bar{V}_A}(e^{-\beta\bar{V}_A} + e^{-\beta\bar{V}_B})^{-1} , \tag{6.4}$$

where $\bar{V}_A$ and $\bar{V}_B$ stand for $\bar{V}(r_A)$ and $\bar{V}(r_B)$, respectively. Thus, by combining Eqs. (6.1), (6.2), (4.7), and (4.8) we obtain

$$\begin{aligned} \lambda &= (r_C - r_A)(2x_Ax_B)^{-1}s(r_C) \\ &\approx \tfrac{1}{2}(e^{-\beta(\bar{V}_A - \bar{V}_B)} + 1)e^{-\beta(\bar{V}_C - \bar{V}_A)} \\ &\approx e^{-\beta A^\dagger} \end{aligned} \tag{6.5}$$

Since $A^\dagger$ is typically much larger than $k_BT, \lambda \ll 1$.

The reader may question whether $N_A(t)$ is not coupled to another slow variable since $N_B(t)$ must also vary with the characteristic time τ_{TS}. But in a closed (canonical ensemble) system, the coupling between $N_A(t)$ and $N_B(t)$ is trivial. The two quantities actually refer to only one dynamical variable since $N_A(t) + N_B(t) = N$ is a constant. The question is not trivial, however, when considering an open (grand canonical ensemble) system. Then diffusion processes make $N(t)$ a dynamical quantity, and $N_A(t)$ and $N_B(t)$ are two truly different variables. Since they are both slow variables, a grand canonical ensemble treatment of isomerization dynamics must treat both variables explicitly if the memory function formalism is to yield a perturbation series ordered with the parameter λ. (See the Appendix for the grand canonical ensemble analysis.)

A *proof* that $N_A(t)$ does not, for a closed system, couple to another slow variable is equivalent to a microscopic derivation of the phenomenological rate law. If other slow variables were involved, then the small s behavior of $\tilde{C}(s)$ would be characterized by more than one relaxation time. We *assume* that the rate law is empirically correct. This assumption leads to two important results. First, since the projected propagator $\exp(QLt)$ involves dynamics orthogonal to N_A, the memory function must relax in a short time, i.e., a time of the order of τ_{mol}. Since $\lambda = \tau_{\mathrm{mol}}/\tau_{\mathrm{TS}} \approx \tau_{\mathrm{mol}}/\tau_{\mathrm{rxn}} \ll 1$, the relaxation time for $M(t)$ is very small compared to τ_{rxn}. Hence, there exists a time Δt that satisfies

$$\tau_{\mathrm{mol}} \ll \Delta t \ll \tau_{\mathrm{rxn}} \approx \lambda^{-1}\tau_{\mathrm{mol}} , \tag{6.6}$$

which can be used with Eq. (5.5) to yield

$$\tau_{\mathrm{rxn}}^{-1} = \int_0^{\Delta t} dt\, M(t)[1 + 0(\lambda)] . \tag{6.7}$$

The second ramification of the assumption uses the identity

$$e^{QLt} = e^{Lt} - \int_0^t dt'\, e^{L(t-t')}PLe^{QLt'} . \tag{6.8}$$

The effect on $-PL$ on a general dynamical variable G can be studied by applying Eqs. (5.3). This gives

$$\begin{aligned} -PLG &= -\delta N_A\langle(\delta N_A)^2\rangle^{-1}\langle\delta N_A LG\rangle \\ &= \delta N_A\langle(\delta N_A)^2\rangle^{-1}\langle\dot{N}_A G\rangle \\ &= -\delta N_A\langle(\delta N_A)^2\rangle^{-1}\langle N\dot{q}_1\delta(r_C - q_1)G\rangle . \end{aligned} \tag{6.9}$$

The second equality follows from time reversal symmetry, and the third equality is obtained from Eq. (4.3). Inspection of Eq. (6.9) shows that PLG is of the order of $\tau_{\mathrm{TS}}^{-1} = \lambda\tau_{\mathrm{mol}}^{-1}$. Thus, provided the projected propagator $\exp(QLt)$ does lead to relaxation in times of the order of τ_{mol}, the integral in Eq. (6.8) is a factor of λ smaller than the unprojected propagator $\exp(Lt)$. Hence, provided N_A is not coupled to another slow variable (which is equivalent to assuming that the experimental rate law is correct), Eqs. (5.2) and (6.8) can be used to write

$$M(t) = \mathfrak{M}(t)[1 + 0(\lambda)] , \tag{6.10}$$

where

$$\mathfrak{M}(t) = \langle(\delta N_A)^2\rangle^{-1}\langle\dot{N}_A e^{Lt}\dot{N}_A\rangle . \tag{6.11}$$

The connection between $M(t)$ and $\mathfrak{M}(t)$ plays a central role in the derivation of a second microscopic formula for τ_{rxn} and in the analysis of transition state theory.

VII. PLATEAU VALUE PROBLEM AND A SECOND CORRELATION FUNCTION EXPRESSION

The function $\mathfrak{M}(t)$ looks like the memory function $M(t)$ except that $\mathfrak{M}(t)$ contains the true propagator $\exp(Lt)$ rather than the projected one. In fact,

$$\mathfrak{M}(t) = -d^2C(t)/dt^2 , \tag{7.1}$$

from which it can be shown that

$$\int_0^\infty dt\, \mathfrak{M}(t) = 0 . \tag{7.2}$$

The Laplace transform of Eq. (7.1) together with Eq. (5.1) yields

$$\tilde{\mathfrak{M}}(s) = s\tilde{M}(s)[s + \tilde{M}(s)]^{-1} , \tag{7.3a}$$

or

$$\tilde{M}(s) = \tilde{\mathfrak{M}}(s)[1 - \tilde{\mathfrak{M}}(s)/s]^{-1} . \tag{7.3b}$$

Equations (7.3) were first discussed by Mori[21] and by Berne *et al.*[23] They can be used to investigate the nature of a plateau in $\mathfrak{M}(t)$ at long times and how appropriate limits can be performed to remove the plateau and obtain the standard flux autocorrelation function expressions for transport coefficients. Recent textbook discussions are given by Forster[15] and by Berne and Pecora.[14]

It turns out that the plateau value is relevant to isomerization dynamics too. The heart of the problem is found in a comparison of Eqs. (5.5), (6.10), and (7.2). According to Eq. (6.10), $M(t)$ and $\mathfrak{M}(t)$ are pointwise identical to within an error of the order of λ. Yet, according to Eqs. (5.5) and (7.2), the integrals over the two functions are completely different. $\tilde{M}(0)$ is τ_{rxn}^{-1}, while $\tilde{\mathfrak{M}}(0)$ is zero. Of course, there is no paradox. This behavior is a consequence of the separation of time scales studied in Sec. VI. In particular, Eqs. (6.7) and (6.10) can be combined to yield

$$\tau_{\mathrm{rxn}}^{-1}=\int_0^{\Delta t} dt\, \mathfrak{M}(t)[1+0(\lambda)]\,. \tag{7.4}$$

In the traditional hydrodynamics applications of Mori–Zwanzig formalism, the dynamical variable under consideration would be a Fourier component of the density of a conserved variable. Then λ would be proportional to k^2, where k is the Fourier wave vector. For a macroscopic system it is possible to take $k\to 0$ and thus $\lambda\to 0$. In that limit, the density becomes the total variable, which is a constant of the motion. The memory function (and the relaxation time) for a constant of the motion is zero. Thus, in the hydrodynamics application one obtains transport coefficients from limits like

$$\lim_{\lambda\to 0}\frac{1}{\lambda}\int_0^{\Delta t} dt\,\mathfrak{M}(t)[1+0(\lambda)]=\int_0^{\infty} dt\, m(t)\,, \tag{7.5a}$$

where

$$m(t)=\lim_{\lambda\to 0}\frac{1}{\lambda}\,\mathfrak{M}(t)\,. \tag{7.5b}$$

The upper limit Δt is replaced by ∞ in the integral due to a scaling relation like Eq. (6.6). [In writing Eq. (7.5b), it is assumed that λ is *defined* so that $\mathfrak{M}(t)$ is proportional to λ to the first power as λ tends to zero. Recall, $M(t)$ and $\mathfrak{M}(t)$ must vanish as $\lambda\to 0$.] While a limit like the one in Eq. (7.5) is meaningful for the hydrodynamic variables, λ is not a *controllable* parameter in the isomerization dynamics problem. Thus, the Δt in Eq. (7.4) cannot be replaced by ∞. Equation (7.4) is applicable only when $A^\dagger$ is large compared to k_BT, so that λ is small. The value of Δt must obey the inequalities in Eq. (6.6).

The importance of the separation of time scales can be appreciated quantitatively by adopting a simple model which contains both τ_{rxn} and τ_{mol}. Suppose

$$\tilde{M}(s)\approx\tau_{\mathrm{TS}}^{-1}+\tau_{\mathrm{mol}}^{-1}(\tau_{\mathrm{rxn}}^{-1}-\tau_{\mathrm{TS}}^{-1})(s+\tau_{\mathrm{mol}}^{-1})^{-1}\,.$$

The leading frequency independent term insures that the short time behavior of $C(t)$ is properly described. The second term imagines that $N_A(t)$ is coupled to molecular processes which relax exponentially on a time scale of τ_{mol}. This second term is zero when transition state theory is exact. The computation of $C(t)$ from this model memory function is straightforward. When λ is small, $C(t)$ is a sum of two exponentials. One relaxes with the time τ_{mol} and guarantees that $C(0+)$ is $-\langle(\delta N_A)^2\rangle\tau_{\mathrm{TS}}^{-1}$. The second and dominant exponential has the relaxation time τ_{rxn}. [If τ_{rxn} is taken to be exactly τ_{TS}, then the short time exponential disappears and the long time one becomes $\langle(\delta N_A)^2\rangle\exp(-t/\tau_{\mathrm{TS}})$.] The area under $\mathfrak{M}(t)=-d^2C(t)/dt^2$ can also be computed for the model. The result is

$$\begin{aligned}\int_0^{\Delta t} dt\,\mathfrak{M}(t)\approx&(\tau_{\mathrm{TS}}^{-1}-\tau_{\mathrm{rxn}}^{-1})e^{-\Delta t/\tau_{\mathrm{mol}}}[1+0(\lambda)]\\&+\tau_{\mathrm{rxn}}^{-1}[1-\Delta t/\tau_{\mathrm{rxn}}+0(\lambda^2)]\,,\end{aligned} \tag{7.6}$$

where we have exhibited the leading correction terms to Eq. (7.4).

Equation (7.4) can be manipulated to yield a physical microscopic formula for the rate constant. By applying Eq. (6.11), Eq. (7.4) becomes

$$\begin{aligned}\tau_{\mathrm{rxn}}^{-1}&=\langle(\delta N_A)^2\rangle^{-1}\int_0^{\Delta t} dt\,\langle\dot{N}_A\dot{N}_A(t)\rangle[1+0(\lambda)]\\&=\langle(\delta N_A)^2\rangle^{-1}\langle\dot{N}_A[N_A(\Delta t)-N_A]\rangle[1+0(\lambda)]\,.\end{aligned} \tag{7.7}$$

In view of Eqs. (4.1)–(4.5), Eq. (7.7) can be expressed as

$$\begin{aligned}\tau_{\mathrm{rxn}}^{-1}&=-(x_Ax_B)^{-1}\langle\dot{q}_1\delta(r_C-q_1)H_A[q_1(\Delta t)]\rangle[1+0(\lambda)]\\&=(x_Ax_B)^{-1}\langle\dot{q}_1\delta(r_C-q_1)H_B[q_1(\Delta t)]\rangle[1+0(\lambda)]\,,\end{aligned} \tag{7.8}$$

where the second equality follows from the first because $H_A(q)=1-H_B(q)$, and the equilibrium ensemble distribution is an even function of $\dot{q}_1$. Recall that $\tau_{\mathrm{rxn}}^{-1}=k_{A\to B}/x_B$, Eq. (2.8). This relationship together with (7.8) yields

$$k_{A\to B}=x_A^{-1}\langle\dot{q}_1\delta(r_C-q_1)H_B[q_1(\Delta t)]\rangle[1+0(\lambda)]\,, \tag{7.9}$$

which is our second correlation function expression for the rate constant.

Equation (7.9) is a principal result of this article. It is also very nearly a standard result. Indeed, the reader should compare Eq. (7.9) with Eq. (17) of Ref. 7 or Eq. (1) of Ref. 6. The most common idea about the approximate nature of Eq. (7.9) is that this equation rests on the correctness of the Boltzmann distribution of the momenta and coordinates for the reactants. The derivation given herein helps clarify this idea. Note that the equilibrium ensemble average appears in our development as a direct consequence of the fluctuation–dissipation theorem. It does not arise from an assumption. However, the separation in time scales $\tau_{\mathrm{mol}}\ll\Delta t\ll\tau_{\mathrm{rxn}}$ draws a picture of the isomerization reaction in which the reactions occur so infrequently (roughly once every τ_{rxn} for each molecule) that the surrounding solvent molecules have sufficient time to conform to and equilibrate with the solute before it reacts again. This equilibrium time is τ_{mol}. The significance of the separation in time scales is surely appreciated in Refs. 8 and 10. However, neither article proceeds far enough to identify the expansion parameter $\lambda=\tau_{\mathrm{mol}}/\tau_{\mathrm{rxn}}$ with the Boltzmann factor for the free energy of activation.

Notice that the ensemble average in Eq. (7.9) can be correctly viewed as the calculation of the average flux across the surface at $q_1=r_C$ given that a reaction is completed successfully at a long time Δt later. The more conventional formula for $k_{A\to B}$ is[6,7]

$$k_{A\to B} = x_A^{-1}\langle q_1\delta(r_C - q_1)\chi_{A\to B}[q_1(\Delta t)]\rangle[1+0(\lambda)]\,, \quad (7.9')$$

where

$$\chi_{A\to B}[q_1(\Delta t)] = H_B[q_1(\Delta t)]H_A[q_1(-\Delta t)] \quad (7.10)$$

is the characteristic function for the $A\to B$ reaction. According to Eq. (7.9′), one averages the flux across the surface at $q_1 = r_C$ given that $q_1(t) > r_C$ at a later time $t = \Delta t$ *and* that $q_1(t) < r_C$ at an earlier time $t = -\Delta t$. It may appear that Eqs. (7.9) and (7.9′) are different. But actually, they are identical. The reason is found by noting that

$$H_A[q_1(-\Delta t)] = 1 - H_B[q_1(-\Delta t)]\,. \quad (7.11)$$

From this equation it is seen that the difference between Eqs. (7.9) and (7.9′) is (within trivial factors) the quantity

$$\langle \dot{q}_1\delta(r_C - q_1)H_B[q_1(\Delta t)]\,H_B[q_1(-\Delta t)\rangle\,,$$

which is the average flux across the surface $q_1 = r_C$ given that $q_1(\Delta t) > r_C$ and $q_1(-\Delta t) > r_C$. According to Liouville's theorem, that averaged flux is zero. Hence, the ensemble averages in Eqs. (7.9) and (7.9′) yield the same numbers.

The interpretation of Eq. (7.9) as an average flux across $q_1 = r_C$ together with Liouville's theorem leads to an important and well-known[6] result: The value of $k_{A\to B}$ computed from the ensemble average in Eq. (7.9) is independent of the location of the surface which divides the reactant, state A, from the product, state B. We have emphasized the necessity of a free energy bottleneck. Its existence leads to the separation in time scales exploited in the analysis of the plateau value and the derivation of Eq. (7.9). But once Eq. (7.9) is established, there is no need to locate r_C at the position of the bottleneck *provided the computation of the ensemble average is computed exactly*. Of course, an approximate evaluation, such as the one performed in transition state theory, can lead to results for $k_{A\to B}$ that depend crucially upon the choice of r_C.

As a final remark we note that in most discussions of Eq. (7.9) the characteristic function $H_B[q_1(\Delta t)]$ is usually described as if Δt is actually ∞. The replacement of Δt with ∞ yields $k_{A\to B} = 0$ for the mathematical reasons stated above. But there is also an important physical reason why the replacement cannot be made. The rate constant $k_{A\to B}$ describes the average frequency for *one* A to B transition per molecule. The right-hand side of Eq. (7.9) clearly fits that physical interpretation provided Δt is larger than τ_{mol} (so that the transition is completed and the degrees of freedom coupled to the reaction coordinate have time to equilibrate) and Δt is smaller than τ_{rxn} (so that back reactions do not clutter our considerations). If Δt gets so large that $\Delta t \gg \tau_{\mathrm{rxn}}$, the molecule will have encountered so many transitions that the characteristic function $H_B[q_1(\Delta t)]$ will be uncorrelated with the initial conditions that placed q_1 at the dividing point r_C. This loss of correlation makes the right-hand side of Eq. (7.9) approach zero since

$$\begin{aligned}&\lim_{\Delta t\to\infty}\langle \dot{q}_1\delta(r_C - q_1)H_B[q_1(\Delta t)]\rangle\\ &\quad = \langle \dot{q}_1\delta(r_C - q_1)\rangle\langle H_B(q_1)\rangle = 0\,.\end{aligned} \quad (7.12)$$

VIII. TRANSITION STATE THEORY

We are now in a position to discuss the validity of classical transition state theory. The characteristic function $H_B[q_1(\Delta t)]$ involves a dynamical calculation. In principle one must compute the time dependence of $q_1(t)$ given the initial condition that $q_1 = r_C$ and all other coordinates and momenta (of solute and solvents) are distributed according to the equilibrium canonical distribution law. If $q_1(\Delta t) > r_C$, the characteristic function is unity. Otherwise it contributes zero to $k_{A\to B}$. To obtain transition state theory we assume

$$\delta(q_1 - r_C)H_B[q_1(\Delta t)] \approx \delta(q_1 - r_C)\theta(\dot{q}_1)\,, \quad (8.1)$$

where $\theta(x)$ is unity for $x > 0$ and zero otherwise. With Eq. (8.1), Eq. (7.9) becomes

$$k_{A\to B} \approx (k_{A\to B})_{\mathrm{TS}} = x_B\tau_{\mathrm{TS}}^{-1}\,. \quad (8.2)$$

According to Eq. (8.1), transition state theory is a consequence of assuming that all trajectories passing the transition point r_C in the reactive direction will indeed be reactive. There will be no recrossings in a time of the order of τ_{mol}. This idea is Wigner's[5] "fundamental assumption."[6] It clearly differs from Eyring's popularized notion that a transition state complex must exist in some sort of quasi-equilibrium with the reactant.[24]

With Eq. (8.1), the rate constant can be calculated from equilibrium statistical mechanics. There is no need to do dynamics. But Eq. (8.1) is clearly a statement about dynamics, and the justification of that approximation requires some dynamical considerations.

Equation (7.9) is independent of the precise location of the transition point r_C. However, Eq. (8.1) and thus transition state theory is sensitive to the choice of r_C. To see why, consider the alternative point r_C' in Fig. 2. With this location for the transition point, $\lambda \ll 1$ so that Eq. (7.9) is still correct. However, due to the negative slopes of both $V(r_C')$ and $\overline{V}(r_C')$, it is unlikely that the requirement of $\dot{q}_1 > 0$ will be sufficient to guarantee that $q_1(\Delta t) > r_C'$ given the initial condition $q_1 = r_C'$. Hence, Eq. (8.1) will not be a good approximation if the transition point is located at r_C'.

To decide upon a good choice for r_C, it is useful to consider the kinetic equation which describes the average motion of q_1 and $\dot{q}_1$. Let

$$F(q,v;t) = \langle\delta[(q - q_1(t)]\delta[v - \dot{q}_1(t)]\rangle_0\,, \quad (8.3)$$

where $\langle\cdots\rangle_0$ denotes the equilibrium ensemble average over initial conditions for the system with $q_1(0)$ and $\dot{q}_1(0)$ fixed at q_0 and v_0. The equation of motion for the derivation of $F(q,v;t)$ from its infinite time value $\delta F(q,v;t) = F(q,v,t) - F(q,v;\infty)$, can be obtained by applying the Mori–Zwanzig formalism. See Chap. 6 of Ref. 15 for a recent textbook discussion. Assume for simplicity that the isomerization process is truly one dimensional. Then, the Mori–Zwanzig theory gives

$$\left(\frac{\partial}{\partial t} - v\frac{\partial}{\partial q} + \mu^{-1}\frac{\partial \bar{V}(q)}{\partial q}\frac{\partial}{\partial v}\right)\delta F(q,v;t) = -\int_0^t dt' \int dq' \int dv'\, m(q,v;q',v';t-t') \times \delta F(q',v';t'), \tag{8.4}$$

where μ denotes the (reduced) mass for the coordinate q_1, and $m(q,v;q',v';t)$ is the memory operator in the kinetic equation. Notice that the potential in the left-hand side, $\bar{V}(q)$, is the potential of mean force for the internal coordinate; that is,

$$\frac{\partial \bar{V}(r)}{\partial r} = -k_B T\frac{\partial \ln s(r)}{\partial r}. \tag{8.5}$$

Hence, the left-hand side of Eq. (8.4) describes the dynamics of the internal coordinates as if the *motion is on a free energy surface*. As a result, a decision about where to place a transition state should be based upon considerations of the *Helmholtz* free energy changes associated with the internal coordinate. For classical isomerization dynamics, it seems that the transition state should be the point r_C—the point of the free energy maximum in Fig. 2.

Another way to decide upon a choice for r_C does not involve any consideration of the dynamics. Rather, one can make use of the inequality

$$[k_{A\to B}]_{\text{TS}} \geq k_{A\to B}. \tag{8.6}$$

This bound for the rate constant is well known and often exploited in statistical approximations for rate constants.[7] It is readily established from Eqs. (8.1), (8.2), and (7.9). Since the transition state theory gives a value for $k_{A\to B}$ that is always greater than the exact value, an optimum choice for r_C is the one which minimizes $[k_{A\to B}]_{\text{TS}}$. From Eqs. (8.2) and (4.7), it is seen that this "optimum" choice corresponds to finding the r_C which minimizes $s(r_C)$. This value of r_C is the location of the maximum to the Helmholtz free energy.

Having fixed r_C, there is still more to be required for transition state theory to be accurate. Indeed, if there is no dissipation, Eq. (8.1) can never work. Suppose the equation of motion for the internal coordinate is conservative. Given the initial conditions $q_1 = r_C$ and $\dot{q} > 0$, a short time later (roughly τ_{mol}), the coordinate will again be at $q_1 = r_C$, but this time with $\dot{q}_1 < 0$. However, the memory operator of the kinetic equation describes dissipation due to the coupling of the reaction coordinate with the bath—the surrounding solvent molecules, and perhaps other degrees of freedom within the molecule itself. If the dissipative effects are large enough, it is unlikely that q_1 will return to r_C in a short time with enough energy to surmount the free energy barrier. Thus, provided dissipation is present and its effects are large enough, transition state theory should be a good approximation. Of course, the bath may also *donate* energy to the isomerizing solute molecule as well as remove it. However, the barrier separating the states is supposed to be high compared to $k_B T$. As a result, the energy associated with the reaction coordinate immediately after it has passed over the barrier will be very high compared to that of a typical degree of freedom in the bath. Hence, the overwhelming circumstance is that after crossing the barrier, the isomer will give energy to the bath rather than accepting it.

It appears that Eq. (8.1) and transition state theory for isomerization dynamics hinges on two conditions: (a) the free energy activation barrier must be high, and (b) the coupling of the bath to the reaction coordinate must be strong enough to dissipate roughly $k_B T$ energy from the isomer in times of the order of τ_{mol}.

To get a quantitative idea for the role of dissipation in isomerization dynamics, imagine that the average bistable potential in Fig. 2 is composed of two square wells separated by a square barrier. Let $A^\dagger$ denote the height of the barrier relative to the minimum of the well at the right. We assume that $A^\dagger \gg k_B T$. Imagine that the "particle" with coordinate q and reduced mass μ has just been passed to the right of the barrier. Its kinetic energy is then very large compared to $k_B T$ so that a simple friction law,

$$\mu\ddot{q} = -\frac{\partial \bar{V}(q)}{\partial q} - \mu\gamma v, \tag{8.7}$$

correctly estimates the average dissipation of energy from the reaction coordinate. In Eq. (8.7), v is the velocity, $\dot{q}$, and γ is the friction constant. Let v_0 denote the velocity of the particle just before it crosses over the edge of the barrier. If v_0 is smaller than some velocity v_m, the frictional dissipation will remove enough energy that the particle will not mount the barrier on its return trajectory (at the far right side of the right square well is an infinite potential wall). That is, trajectories with $v_0 < v_m$ will be trapped in the right square well. However, if $v_0 > v_m$, the particle will recross the edge of the barrier, and Eq. (8.1) will probably be violated. The equation for v_m is

$$v_m^2 = (2w\gamma)^2 + 4w\gamma(2A^\dagger/\mu)^{1/2}, \tag{8.8}$$

where w is the width of the right square well. In view of Eq. (8.1) and (7.9) and definition of v_m, we are led to estimate the fractional error due to Eq. (8.1) as the ratio

$$R = \int_{v_m}^{\infty} v_0 \exp(-\beta v_0^2\mu/2)dv_0 \Big/ \int_0^{\infty} v_0\exp(-\beta v_0^2\mu/2)dv_0 = \exp\{-2\beta\mu w\gamma[(w\gamma) + (2A^\dagger/\mu)^{1/2}]\}. \tag{8.9}$$

For a liquid solvent $w\gamma \approx w\tau_{\text{mol}}^{-1} \approx \langle|\dot{q}|\rangle$. The typical size of a barrier to rotation about a single bond is 3 kcal/mole or more. By using these estimates, one finds $R \lesssim 0.01$. It would appear that Eq. (8.1) and thus transition state theory is a good approximation for isomerization dynamics in liquids. However, the dissipation which makes R so small gives rise to another mechanism which can make Eq. (8.1) a poor approximation.

The calculation outlined by the previous paragraph is applicable after the particle has reached the edge of the barrier. But is must first travel there from the midpoint of the barrier without recrossing that midpoint. (We assume the transition state is located at the midpoint.) During this trajectory, the velocity of the particle is not necessarily large, and as a result the medium can alter the velocity in ways that are more complicated than those resulting from frictional dissipation. If the barrier is wide enough, the medium will have an excellent opportunity to reverse the trajectory of the particle. This picture of diffusional motion at the top of a barrier is at the heart of Kramers'[25] theory which employs the Fokker–Planck equation to describe the effects of fluctuating forces due to the medium. We present now a simple collision theory treatment of this phenomenon.

Picture the motion on top of the barrier as a random walk with step length $v\tau_{\text{coll}}$, where v denotes the initial step velocity of the particle and τ_{coll}^{-1} is the average frequency at which the medium strikes the particle. Take a simplified random walk model in which the particle has a 50–50 chance of getting to either edge of the barrier independent of the initial step direction unless the initial step is larger than half the barrier width. If the step length is larger than the half width, there is unit probability that the particle will reach the right edge if the initial velocity (i.e., step) is to the right, and unit probability to reach the left edge if the initial velocity is to the left. We have seen that once the particle is at the edge, it is (to an excellent approximation) always trapped in the adjacent well by dissipation. Hence, according to this random walk model and Eqs. (7.9) and (8.1), the transmission coefficient, $k_{A\to B}[k_{A\to B}]_{\text{TS}}^{-1}$, is the ratio

$$R' = \left(\int_0^\infty v\exp(-\beta\mu v^2/2)dv\right)^{-1}\left(\frac{1}{2}\int_{-L/\tau_{\text{coll}}}^{0} v\exp(-\beta\mu v^2/2)dv + \frac{1}{2}\int_0^{L/\tau_{\text{coll}}} v\exp(-\beta\mu v^2/2)dv + \int_{L/\tau_{\text{coll}}}^{\infty} v\exp(-\beta\mu v^2/2)dv\right)$$

$$= \exp(-\beta\mu L^2/2\tau_{\text{coll}}^2) = \exp(-L^2/2\tau_{\text{coll}}^2\,\langle|\dot{q}|^2\rangle)\,, \qquad (8.10)$$

where L is the half width of the barrier.

According to Eq. (8.10), transition state theory ($R' = 1$) will incur a significant error if the average velocity of the reaction coordinate is not larger than L/τ_{coll}. Since the collision frequency τ_{coll}^{-1} can be very large in a liquid it is quite likely that for a large number of isomerization reactions

$$L/\tau_{\text{coll}} \approx \langle|\dot{q}|\rangle\,.$$

Furthermore, since τ_{coll} is a strong function of the liquid density, it is not correct to interpret the transmission coefficient R' as a simple multiplicative constant. Thus, our estimates indicate that it is worthwhile to pursue the development of systematic molecular theory for isomerization dynamics.

ACKNOWLEDGMENTS

While carrying out the research reported in this article, I have benefitted from conversations with Bruce J. Berne, Philip Pechukas, Lawrence R. Pratt, and Robert Zwanzig.

APPENDIX A: ENSEMBLE DEPENDENCE OF $\langle(\delta N_A)^2\rangle$

For an open system containing interconverting A and B solute isomers at low dilution, the average number fluctuations obey the equations

$$\langle(\delta N_A)^2\rangle = \langle N_A\rangle$$

and

$$\langle(\delta N_B)^2\rangle = \langle N_B\rangle\,. \qquad (A1)$$

However, if the system is closed, the correct result is the one described in Hill's textbook,[19]

$$\langle(\delta N_A)^2\rangle = \langle(\delta N_B)^2\rangle = \langle N_A\rangle\langle N_B\rangle/N\,. \qquad (A2)$$

In the canonical ensemble, N_A and N_B are coupled due to the constraint $N = \text{constant} = N_A + N_B$. Hence, $\delta N_A = -\delta N_B$. In this Appendix we find that this coupling leads to a long range tail in molecular pair correlations which is absent in the grand canonical ensemble. This difference in the long range correlations is responsible for the difference between Eq. (A2) (the canonical ensemble result) and Eq. (A1) (the grand canonical ensemble result).

We adopt the general notation used in the Chandler–Pratt theory of chemical equillibria.[1,2] For simplicity, consider only the solute correlations so that we can omit subscripts which identify the molecular species. According to Eq. (4.33) of Ref. 2,

$$\chi(1,2) = \langle\delta\rho(1)\delta\rho(2)\rangle$$
$$= \rho s(1)\delta(1-2) + 0(\rho^2)\,. \qquad (A3)$$

Note that 1 and 2 are abbreviations for all the coordinates necessary to specify the locations and conformations of solute molecules 1 and 2, respectively. From Eq. (A3), we find at low solute concentrations

$$\begin{aligned}\langle(\delta N_A)^2\rangle &= \int d1\int d2\,H_A(1)H_A(2)\chi(1,2)\\ &= \int d1\int d2\,H_A(1)H_A(2)\rho s(1)\delta(1-2)\\ &= \int d1\,H_A(1)H_A(1)\rho s(1)\\ &= \int d1\,H_A(1)\rho s(1)\\ &= \langle N_A\rangle\,, \end{aligned} \qquad (A4)$$

which is a derivation of Eq. (A1). It is a grand canonical ensemble result because Eq. (A3) is correct for that ensemble only.

To obtain the canonical ensemble result, a scheme similar to that of Lebowitz and Percus[12] can be used. Let $\rho_N^{(2)}(1,2)$ denote the pair distribution for isomers 1 and 2 in a closed system with N isomers. The grand canonical $\rho^{(2)}(1,2)$ is obtained from $\rho_N^{(2)}(1,2)$ by averaging N. Since the distribution for N will be sharply peaked (for a macroscopic system) near $\langle N\rangle$, we can expand $\rho_N^{(2)}(1,2)$ about $\rho_{\langle N\rangle}^{(2)}(1,2)$, and then average the Taylor expanded function. By retaining terms through quadratic order, one finds

$$\rho^{(2)}(1,2) = \rho_{\langle N\rangle}^{(2)}(1,2) + \tfrac{1}{2}\langle(\delta N)^2\rangle\frac{\partial^2}{\partial\langle N\rangle^2}\,\rho_{\langle N\rangle}^{(2)}(1,2)\,. \qquad (A5)$$

For a macroscopic system, the second term on the right-hand side is vanishingly small at each pair of configurations 1 and 2. However, we are interested in an integral of $\rho^{(2)}(1,2)$ over all space. For this purpose, it is necessary to retain asymptotic terms. At low solute concentrations, Eq. (A5) gives

$$\rho^{(2)}(1,2)=\rho^{(2)}_{\langle N\rangle}(1,2)+\langle N\rangle^{-1}\rho^2 s(1)s(2). \tag{A6}$$

Hence, the grand canonical and canonical $\chi(1,2)$'s are related by

$$\chi_{\langle N\rangle}(1,2)=\chi(1,2)-\rho^2 s(1)s(2)/\langle N\rangle. \tag{A7}$$

To obtain the canonical $\langle(\delta N_A)^2\rangle$, multiply $\chi_{\langle N\rangle}(1,2)$ by $H_A(1)$ and $H_A(2)$ and integrate. This gives

$$\begin{aligned}\langle(\delta N_A)^2\rangle_{\langle N\rangle}&=\langle(\delta N_A)^2\rangle-\langle N\rangle^{-1}\left(\int d1\,\rho s(1)\,H_A(1)\right)^2\\&=\langle N_A\rangle-\langle N\rangle^{-1}\langle N_A\rangle^2\\&=\langle N_A\rangle[1-\langle N_A\rangle/(\langle N_A\rangle+\langle N_B\rangle)]\\&=\langle N_A\rangle\langle N_B\rangle/\langle N\rangle,\end{aligned} \tag{A8}$$

which is a derivation of Eq. (A2).

APPENDIX B: DERIVATION OF CORRELATION FUNCTION FORMULAS FOR $k_{A\to B}$ FOR AN OPEN SYSTEM

The analysis for an open system must begin with a rate law that is more complicated than Eq. (2.4). The time dependence of $N_A(t)$ and $N_B(t)$ now depends upon diffusion in and out of the system as well as upon interconversion. To account for this additional mechanism, we write

$$\dot n_A(k,t)=-k^2D_An_A(k,t)-k_{A\to B}n_A(k,t)+k_{B\to A}n_B(k,t), \tag{B1a}$$

and

$$\dot n_B(k,t)=-k^2D_Bn_B(k,t)-k_{B\to A}n_B(k,t)+k_{A\to B}n_A(k,t), \tag{B1b}$$

where $n_A(k,t)$ denotes the nonequilibrium average of

$$\rho_A(k,t)=\sum_{j=1}^{N(t)}H_A[q_j(t)]e^{-i\mathbf{k}\cdot\mathbf{R}_j(t)}, \tag{B2}$$

and $n_B(k,t)$ has a similar definition. The vector $\mathbf{R}_j(t)$ is the position of the center of the jth solute molecule at time t. There are $N(t)$ such molecules in the system at that time. The diffusion constants describing the rate of flow of the A and B solute isomers are D_A and D_B, respectively.

The phenomenological equations (B1) can be used in conjunction with the fluctuation–dissipation theorem to describe the small k and s behavior of

$$\tilde C_{ij}(k,s)=\int_0^\infty dt\,e^{-st}\langle\delta\rho_i^*(k,0)\delta\rho_j(k,t)\rangle. \tag{B3}$$

That procedure yields

$$\tilde C_{ij}(k,s)\sim\{[s\mathbf{1}+\mathbf{L}(k)]^{-1}\}_{ji}\langle N_i\rangle, \tag{B4}$$

for small k and s, where $\mathbf{1}$ is the identity matrix, $[\ ]^{-1}$ denotes the inverse of the matrix, and $\mathbf{L}(k)$ has elements

$$\begin{aligned}L_{AA}(k)&=k^2D_A+k_{A\to B},\\L_{AB}(k)&=-k_{B\to A},\\L_{BA}(k)&=-k_{A\to B},\\L_{BB}(k)&=k^2D_B+k_{B\to A}.\end{aligned} \tag{B5}$$

To obtain Eq. (B4), one must note that for an open system

$$\langle\delta N_i\delta N_j\rangle=\delta_{ij}\langle N_i\rangle, \tag{B6}$$

which follows from the analysis sketched in Appendix A. After some matrix algebra, Eq. (B4) yields

$$\lim_{k\to0}[\tilde C_{AA}(k,s)-\tilde C_{AB}(k,s)]\sim\langle N_A\rangle\,[s+\tau_{\mathrm{rxn}}^{-1}]^{-1}, \tag{B7}$$

for small s. Hence, in the grand canonical ensemble

$$\lim_{s\to0}\int_0^\infty dt\,e^{-st}\langle\delta N_A[\delta N_A(t)-\delta N_B(t)]\rangle=\langle N_A\rangle\tau_{\mathrm{rxn}}, \tag{B8}$$

where we have noted that

$$\lim_{k\to0}\delta\rho_i(k,t)=\delta N_i(t). \tag{B9}$$

The reader should compare Eq. (B8) with the canonical ensemble formula Eq. (3.6). The two equations look different.

The memory function equation for τ_{rxn} analogous to Eq. (5.5) is obtained by applying the Mori–Zwanzig formalism to the *pair* of dynamical variables, $\rho_A(k,t)$ and $\rho_B(k,t)$. [See the remarks following Eq. (6.5)]. This procedure introduces a memory function matrix given by the usual formulas.[20–22] For example, with Eq. (B6) one finds

$$M_{AA}(k,t)=\langle N_A\rangle^{-1}\langle\dot\rho_A^*(k,t)e^{QLt}\dot\rho_A(k,t)\rangle, \tag{B10}$$

where $Q-1=P$, and P projects onto the vector $[\delta\rho_A(k),\delta\rho_B(k)]$. The memory function matrix formula for $\tilde C_{ij}(k,s)$ is

$$\tilde C_{ij}(k,s)=\{[s\mathbf{1}+\tilde{\mathbf{M}}(k,s)]^{-1}\}_{ij}\langle N_i\rangle, \tag{B11}$$

which can be compared with Eqs. (B4) and (B5) to yield

$$k_{A\to B}=\lim_{s\to0}\lim_{k\to0}\tilde M_{AA}(k,s). \tag{B12}$$

Notice that this formula looks different than the canonical result, Eq. (5.5).

The arguments about time scales discussed in Secs. VI and VII can now be applied to $\tilde M_{AA}(k,s)$. Those arguments combined with Eqs. (B11) and (B12) yield

$$\begin{aligned}k_{A\to B}&=\lim_{s\to0}\lim_{k\to0}\int_0^\infty dt\,e^{-st}\langle N_A\rangle^{-1}\left\langle\left(\sum_{i=1}^N-q_i\delta(r_C-q_i)\right)\right.\\&\qquad\left.\times e^{QLt}\frac{d}{dt}\left(\sum_{j=1}^N H_A[q_j(t)]\right)\right\rangle\\&=\frac{\langle N\rangle}{\langle N_A\rangle}\langle-\dot q_1\delta(r_C-q_1)H_A[q(\Delta t)]\rangle[1+0(\lambda)]\\&=x_A^{-1}\langle\dot q_1\delta(r_C-q_1)H_B[q(\Delta t)]\rangle[1+0(\lambda)].\end{aligned} \tag{B13}$$

This final physical equation is identical to the rate constant formula, Eq. (7.9), derived in the main text by employing the canonical ensemble.

[1] D. Chandler and L. R Pratt, J. Chem. Phys. 65, 2925 (1976).

[2] L. R. Pratt and D. Chandler, J. Chem. Phys. 66, 147 (1977).

[3] P. J. Flory, *Statistical Mechanics of Chain Molecules* (Interscience, New York, 1969).

[4] L. R. Pratt, Ph.D. thesis, (University of Illinois, 1977).

[5] E. Wigner, Trans. Faraday Soc. 34, 29 (1938).

[6] W. H. Miller, Acc. Chem. Res. 9, 306 (1976).

[7] P. Pechukas in *Dynamics of Molecular Collisions; Part* B, edited by W. H. Miller (Plenum, New York, 1976).

[8] T. Yamamoto, J. Chem. Phys. 33, 281 (1960).

[9] R. Zwanzig, Ann. Rev. Phys. Chem. 16, 67 (1965).

[10] H. D. Kutz, I. Oppenheim, and A. Ben-Reuven, J. Chem. Phys. 61, 3313 (1974).

[11] F. H. Stillinger, *Advances in Theoretical Chemistry* , 3, edited by H. Eyring and D. Henderson, (in press).

[12] J. L. Lebowitz and J. K. Percus, Phys. Rev. 122, 1675 (1961).

[13] M. S. Green, Phys. Rev. 119, 829 (1960).

[14] B. J. Berne and R. Pecora, *Dynamic Light Scattering* (Interscience, New York, 1976); Sec. 11.6 and 11.B.

[15] D. Forster, *Hydrodynamic Fluctuations, Broken Symmetry, and Correlation Functions* (Benjamin, Reading, Ma. 1975) Sec. 5.4.

[16] R. Zwanzig, J. Chem. Phys. 40, 2527 (1964).

[17] P. C. Martin, in *Statistical Mechanics of Equilibrium and Non-Equilibrium*, edited by J. Meixner (North-Holland, Amsterdam, 1965).

[18] D. Forster, Ref. 15, Chap. 2.

[19] T. L. Hill, *Introduction to Statistical Thermodynamics* (Addison-Wesley, Reading, Ma, 1960) Sec. 10.3.

[20] R. Zwanzig, in *Lectures in Theoretical Physics, Vol.* 3 (Interscience, New York, 1961).

[21] H. Mori, Prog. Theor. Phys. (Kyoto) 34, 423 (1965).

[22] D. Forster, Ref. 15 Chap. 5; B. J. Berne and R. Pecora, Ref. 14 Chap. 11.

[23] B. J. Berne, J. P. Boon, and S. A. Rice, J. Chem. Phys. 45, 1086 (1966).

[24] K. J. Laidler, *Chemical Kinetics* (McGraw-Hill, New York, 1965), Chap. 3.

[25] H. A. Kramers, Physika 7, 284 (1940).